

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: An, Gang
O'Hara, S. Mark
Ralph, David
Veltri, Robert

(ii) TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE

(iii) NUMBER OF SEQUENCES: 87

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Arnold, White & Durkee
(B) STREET: P.O. Box 4433
(C) CITY: Houston
(D) STATE: Texas
(E) COUNTRY: USA
(F) ZIP: 77210

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US Unknown
(B) FILING DATE: Concurrently Herewith
(C) CLASSIFICATION: Unknown

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/692,787
(B) FILING DATE: 31-JUL-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Nakashima, Richard A.
(B) REGISTRATION NUMBER: P-42,023
(C) REFERENCE/DOCKET NUMBER: UROC:018

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (512) 418-3000
(B) TELEFAX: (512) 474-7577

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 391 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCCAGTCGC TCAGAAATTT CCTTTGATGC TTTGAAGTTA TCTCTCTTGG ATCTGCTTCC	60
TCCTTATCGT CTCTACATCC CAAGAACAGA GAGTGAGTCT TCTTTATTTT CTTATCTCTG	120
TTTTTAGCAC AGTATTTGAT ATATAGTGTA GATACTATAA ATGCTTGCTA AACTTTGTCA	180
AATTCACAT TTTTAAAATA AAAATGAGAA TGAGCTTGTA GTCAACATGG CGTTTGTAAG	240
TTTGGAGTCT ATATATGGTA GATATACATA TTTTTAAATC TAAGTGCAAC TTTTCTCTTG	300
ATTATCTTGA AATGCCTTAT CATCTCCACA TTTGCTGTAG GCAGTAGTTT AGTGGGTCCA	360
TTATATCTGC CAACTGATT GTCTTAAATA A	391

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 614 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAGTAGTGGC CCCAAATGCC AGGCTGCACT GATATTTATT GGATATAAGA CAAAGGGGCA	60
GGGTAAGGAA TGTGAACCAT CTCCAATAAT AGGTAAGGTC ACATGGGTCA TGTGTCCACT	120
GGACAGGGGG CCCTTCCTG CCTGGCAGCA GAGGCAGAGA GAGAGAGAAG AGAGAGAGAC	180
AGCTTATGCC ATTATTTCTG CATATCAGAC ATTTAGTACT TTCACTAATT TGCTCCTGCT	240
ATCTAAAAGG CAGAGCCAGG TATACAGGAT GGAACATGAA AGCGGACTAG GAGCGTGACC	300
ACTGAAGCAC AGCATCACAG GGAGACAGGC CTCTGGATAC TGGCCGGGGG GCCCTGACTG	360
ATGTCAAGGC CCTCCACAAG AGTGGAGGAG TTAGTCTTCC TCTAAACTCC CCCGGGGGAA	420
AGGGAGGCTC CTTTTCCAG TCTGCTAAGT AGTGGGTGTT TTTCTTGAC ACTGATGCTA	480
CTGCTAGACC ATGGTCCACT TTGCAACAGG CATCTTCCCA GAACTGGTG TTAGTGCTAG	540
ACCAAGCCCT CTGGTGGCCC TGTCCGGGCA TAAGAGAAGG CTCACACTCT TGTCTTCTGG	600
CCACTTCGCA CTAT	614

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 757 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACAACGACAC ATTCAGGAGT TAAATATTTA TCATCAAACA TTGGATTTTT CCTTAACGCT	60
AGAGATTGCT ACAAATCTTC TGAAGGGTCT CAATGGCTTC AGGCTAAGAA GAGATTCTC	120
CCTGTTATAA GCAGCAAGAC AAATTAGCCA TTCTACTCTC AAACCTCACT AATGATCACA	180
TTCTTTCCAA AAGGAACTCT AGAAGACCAA ATGCCCCGAG TTAAGAACAT CAAAACCTAAC	240
CATCTGAAGA AACTTCCCAA GTGTAAGACT CTGCCATTAA AACATTACCG AGAGGGGACT	300
CAAACAGTCT TTTCTTCCCT TTGTCGTGTT TCTTTGCTCC CAGACCCAAG GCACTTGGCG	360
GACAGTACTT GATACAATAA TTTAAAAAGC ACCACTCCCT TCCCACTTTG TAAATACCCA	420
GAACCTAAT TGGACCACCC TGAAGCTTAG GACCTACCAG CCATACAAAT AGTAAACTCT	480
GTCCACGATT CACTCATCTG TGTATTTTCT ATAGATGTTT ACTAGGCGTT TGTTATATAA	540
AAATACCCCG GCCAGGCACG GTGGCTCACG CCTGTAATCC CAGCACTTTG GGAGGTGGGT	600
GGATCACCTG AGGTCGGGAG TTCGAGACCA GCCTGACCAG CATGGTGGAA CCCCATCTC	660
TACTAAAAAC ACAAAAAATT AGCCGGGCGT GGTGGCACAT GCCTGTAATC CCAGCTACTC	720
AGGAGGCTGA GCGGAGAAT TGCTTGAACC CGGAAGG	757

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 673 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CAGGACACAG AGTAAGATAC CCACTGACTT CTTGTGGTCT ACTTCCTGGG TGTTGTTTCA	60
ATGGGCTTTG TTATAACAGG ACTAGTCTTC TGTAATACA ACTTGGTAAA TAGGATGAAA	120
CATAACTTTG CGACAATTCA GTAGAAATAG GCATACAAAC CTGGGCCTGA TGACACTCAC	180
CTCCCCTTGG CTATAAACAT TACCCTACCT GTTAAGTCAG TAATCCTTTG GGAGAGCGCT	240
TACTGAGTAT CTATGATATG CAAAGACCAA AGACCGAGGG GGATCCCTGG TGTAGAGCAA	300

GCACACACCT GGTATTAGC TACCTGCCAC CCTGCTGGGC ATGCAACATA CATTGTCTCA	360
AATTCTAACC ACCCTGCAAG GCAAGCTTCC TTGTTCTTTT AAAGAAGAAA AGTAGACCAG	420
CAAGATTGAT TTGCTCAAGA TTACACAGCC TGGAACTCTG TCATGGGCAT GTCTGACTCT	480
GATAGCAATA CCCTCAAAGA AACTGTCAGA GAAGACTCAA TAAGAAGAAA GTTGAGATAC	540
AGAAACCAAC AGGAGAAGGT AATTCAGAAA TTCAAACAGA GTGGGTGTGA TGGGAAGAAT	600
TCATTAATAA GAAGGTACCT CTGTAGAAAA ATCTTACCAG ACAGTCTGGA AGTGAAGGAA	660
ACAGCCAATA GTC	673

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 358 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCACTGCAC ATTAAGATGG AGCCCGAAGA GCCCACTCC GAGGGGGCAT CGCAGGAGGA	60
TGGGGCTCAA GGTGCCTGGG GCTGGGCACC CCTAAGTCAC GGCTCTAAGG AGAAAGCTCT	120
CTTCCTGCCC GCGGAGCCC TCCCCTCCCC CCGGATCCCC GTGCTTTCCC GAGAGGGGAG	180
GACCAGAGAC CGGCAGATGG CTGCAGCGCT CCTCACTGCC TGGTCCCAGA TGCCAGTGAC	240
TTTCGAGGAT GTGGCCTTGT ACCTCTCCCG GGAGGAGTGG GGACGGCTGG ACCACACGCA	300
GCAGAACTTC TACAGGGAAT GTCCTGCAGA AGAAAAATGG GCTGTCACTG GGCTTTCC	358

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CACAGATGTA GCTTCCTCAC TGG	23
---------------------------	----

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:

SECRET

CTGGAGTACA	ATGTCAGTGT	TTACACTGTC	AAGGATGACA	AGGAAAGTGT	CCCTATCTCT	60
GATACCATCA	TCCCAGCTGT	TCCTCCCTCCC	ACTGACCTGC	GATTCAACCA	CATTGGGTCCA	120
GACACCATGC	GTGTCACCTG	GGCTCCACCC	CCATCCATTG	ATTTAACCA	CTTCCTGGTG	180
CGTTACTCAC	CTGTGAAAAA	TGAGGAAGAT	GTTGCAGAGT	TGTCAATTTT	TCCTTCAGAC	240
AATGCAGTGG	TCTTAACAAA	TCTCCTGCCT	GGTACAGAAT	ATGTAGTGAG	TGTCTCCAGT	300
GTCTACGAAC	AACATGAGAG	CACACCTCTT	AGAGGAAGAC	AGAAAACAGG	TCTTGATTCC	360
CCAACTGGCA	TTGACTTTTC	TGATATTACT	GCCAACTCTT	TTACTGTGCA	CTGGATTGCT	420
CCTCGAGCCA	CCATCACTGG	CTACAGGATC	CGCCATCATC	CCGAGCACTT	CAGTGGGAGA	480
CCTCGAGAAG	ATCGGGTGCC	CCACTCTCGG	AATTCCATCA	CCCTCACCA	CCTCACTCCA	540
GGCACAGAGT	ATGTGGTCTCAG	CATCGTTGCT	CTTAATGGCA	GAGAGGAAAG	TCCCTTATTG	600
ATTGGCCAAC						610

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1649 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

CGGCAGCCAG	CCTATTCCTTT	GGCCGGGTCG	GTGCGAGTGG	TCGGCTGGGC	AGAGTGCACG	60
CTGCTTGCGC	CCGCAGGTGA	TCCCGCCGTC	CAC'TCCCGGG	AGCAGTGATG	TTGGGCAACT	120
CTGCGCCGGG	GCCTGCGACC	CGCGAGGCGG	GCTCGGCGCT	GCTAGCATTG	CAGCAGACGG	180
CGCTCCAAGA	GGACCAGGAG	AATATCAACC	CGGAAAAGGC	AGCGCCCGTC	CAACAACCGC	240
GGACCCGGGC	CGCGCTGGCG	GTACTGAAGT	CCGGGAACCC	GCGGGGTCTA	GCGCAGCAGC	300
AGAGGCCGAA	GACGAGACGG	GTTGCACCCC	TTAAGGATCT	TCCTGTAAAT	GATGAGCATG	360
TCACCGTTCC	TCCTTGAAAA	GCAAACAGTA	AACAGCCTGC	GTTCAACATT	CATGTGGATG	420

AAGCAGAAAA	AGAAGCTCAG	AAGAAGCCAG	CTGAATCTCA	AAAAATAGAG	CGTGAAGATG	480
CCCTGGCTTT	TAATTCAGCC	ATTAGTTTAC	CTGGACCCAG	AAAACCATTG	GTCCCTCTTG	540
ATTATCCAAT	GGATGGTAGT	TTTGAGTCAC	CACATACTAT	GGACATGTCA	ATTGTATTAG	600
AAGATGAAAA	GCCAGTGAGT	GTTAATGAAG	TACCAGACTA	CCATGAGGAT	ATTCACACAT	660
ACCTTAGGGA	AATGGAGGTT	AAATGTAAAC	CTAAAGTGGG	TTACATGAAG	AAACAGCCAG	720
ACATCACTAA	CAGTATGAGA	GCTATCCTCG	TGGACTGGTT	AGTTGAAGTA	GGAGAAGAAT	780
ATAAACTACA	GAATGAGACC	CTGCATTTGG	CTGTGAACTA	CATTGATAGG	TTCCTGTCTT	840
CCATGTCAGT	GCTGAGAGGA	AAACTTCAGC	TTGTGGGCAC	TGCTGCTATG	CTGTTAGCCT	900
CAAAGTTTGA	AGAAATATAC	CCCCCAGAAG	TAGCAGAGTT	TGTGTACATT	ACAGATGATA	960
CCTACACCAA	GAAACAAGTT	CTGAGAATGG	AGCATCTAGT	TTTGAAAGTC	CTTACTTTTG	1020
ACTTAGCTGC	TCCAACAGTA	AATCAGTTTC	TTACCCAATA	CTTTCTGCAT	CAGCAGCCTG	1080
CAAAC TGCAA	AGTTGAAAGT	TTAGCAATGT	TTTTGGGAGA	ATTAAGTTTG	ATAGATGCTG	1140
ACCCATACCT	CAAGTATTTG	CCATCAGTTA	TTGCTGGAGC	TGCCTTTTCAT	TTAGCACTCT	1200
ACACAGTCAC	GGGACAAAGC	TGGCCTGAAT	CATTAATACG	AAAGACTGGA	TATACCCTGG	1260
AAAGTCTTAA	GCCTTGCTCT	ATGGACCTTC	ACCAGACCTA	CCTCAAAGCA	CCACAGCATG	1320
CACAACAGTC	AATAAGAGAA	AAGTACAAAA	ATTCAAAGTA	TCATGGTGTT	TCTCTCCTCA	1380
ACCCACCAGA	GACACTAAAT	CTGTAACAAT	GAAAGACTGC	CTTTGTTTTTC	TAAGATGTAA	1440
ATCACTCAAA	GTATATGGTG	TACAGTTTTT	AACCTTAGGTT	TTTAATTTTA	CAATCATTTT	1500
TGAATACAGA	AGTTGTGGCC	AAGTACAAAT	TATGGTATCT	ATTACTTTTT	AAATGGTTTT	1560
AATTTGTATA	TCTTTTGTAT	ATGTATCTGT	CTTAGATATT	TGGCTAATTT	TAAGTGGTTT	1620
TGTTAAAGTA	TTAATGATGC	CAGCTGCCG				1649

(2) INFORMATION FOR SEO ID NO:9:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTCTGGGGCT CTTACTATAA AAGGGGACCA ACTCTCCCTT TGTCATATCT TGTTTCTGAT 120
GACAAAAAAT AACACATTGT TAAAATTGTA AAATTAAAAC ATGAAATATA AATTA 175

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTTTCGCTCC ACATTCATCC TTTCTTACTG GGCAGTGATG TTGAGAGCAT CAGGCAGGGT 60
ATAATGTTAT GTTGACAGTAA CAAACACCCT CAATATCTCA GTGGCTTAAA ATGACAACGA 120
TCTTTTTTTT GTTTGTTTGT TTATGCTCTA TATCACCCAG GGATCA 166

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGCTCTGCCC CACATCTGAA CAAGCTAATA AGAAAGCCCG ATGTTCTTTC CTTTGGTGCC 60
ATTGGGAAAT TCAAACCATG CACAACTCTG CCTGTATGAA GGGCGCA 107

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 183 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CAACCTTAGC CCCTCTCCTC TTCTTCACGA TGCCATTCTG CCATTTCTGT TTTGTGGTAG 60
ACAGGTTGGC CCAGGCACTC TAAGGCCAG GCTGGCACAG GTTGGCCCAG GCACTTCAAG 120
CCTAAGTCCA TTTACAGTTT CTATTCCATC TCTTCCTAAA GAAGAGGAGA GGGGCTAAGG 180

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 92 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAACAAACGT CTTTGGGTAA AATTCTATTT CTTTAAATGT TTTAAAATAT TTGTAGTCAC 60
 TAATTGTAAG TCATATTCCT CTTTGTCCAG CT 92

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GATGTAATTA AAGCTGTAGA TGAGGGCTAT CGACTGCCAC CCCCCATGGA CTGCCCAGCT 60
 GCCTTGATATC AGCTGATGCT GGACTGCTGG CAGAAAGACA GGAACAACAG ACCCAAGTTT 120
 GAGCAGATTG TTAGTATTCT GGACAAGCTT ATCCGGAATC CCGGCAGCCT GAAGGATCAT 180
 CA 182

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCCAAATGGG TAGCATTGTT GCTCGGCCTT CTAGTCTGCC AGTAGGAAAG TCCAACCATT 60
 AGGTCGGGGA AGAAGGGTCT GGATTTGGTT GACAATGGTT GGATGGGGGA TAGAAGCAGA 120
 GAGAGAGAGG GAGGGCAGCT CAAGGGTATC TTGCCCCACT CTGTTTATGC TGAT 174

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CACCTAACAA TATATCAATT TTTTAAAAAT GGAATTTCTT ATGCCCTCTT TATTTATGGA 60
CATGTATGTC CATAATGGGA GACGTTTCTT TTGGACTGAT GCTTGAATCA GTGGGTGCTT 120
GGCATTGCTG AT 132

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CAGACACACA CATGCACACC ATTCTAGAAT GCTTCCTTAA AAGAAGGAGG GTTGCCCTAG 60
TCTCAAAATC TTAAAAGCCA TATGTGCATT GATTTCTGCA CAGGTAGGCA ATTTGTGATT 120
TTATTTTTTCC TTATG 135

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTTCATGGCA GGACTCGGTT TGGG 24

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 471 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCCCCAAATG CCAGGCTGCA CTGATCTCAT GTCTGTGTCA CTGGAACCAA CAGGCCTGCC 60
TCAACCACTG TCCACCTGCA CATCTGAGAG GCTGGCAGGT CACCAGGGCT AGCCGTGCAC 120
GTCAGTTCCT GGAAGAAAAG TAGAATGTGA ATCATCTTCT CTCAAACGCC TATCAAAAGC 180
CCAGCTGAGA TCAATAATTT GGTGGGAGAA CAGACCTGTA CCAATTGGCT CGGTGTTTGG 240
TGGGGTATTG TAAATTTGGA TCCTAAATCA AAGGGTATCC CTAGAAGGAC CCACATGGAA 300
TGGCCTCCTC CTAAACATCC CTCCATGTTG GTACTTCCTG ACTCTTTTCC AGCAATCTCA 360
AAGCACAAGA AGCAGTGGTG GGAACCCAGG CCTGGCATCT TGTGGAGCC CATGGTTGGG 420
GGGTAGGAGC AACTTTACAG GCCATCAATT ATGCCCCTAT ACGCACCTCC C 471

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 209 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCCCTTTATA AATACGATTA GTATGGAGAA TTGATACATT AACAGTTAGC TTTATAAAATT 60
GACAGATTTTC TAAATTAACC TATGGTCCAC AAATCAAGTT CTATCACTAT TTCCTGCCAC 120
CAAAATCAGT GATGAAGCCT CTCCACACT AAATGAAGAG TGGCGAGGGA CAGAATTCCA 180
CTTGTCTTCC TTTTGCTGCA CTAATACTACA 209

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 407 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CAAGCAGCAT AGCCTCTCTG AACTCAATT TCCTCACATT TATAAATGAG CTTTATATT 60
ATTTACAAAC CTACCTCATA GAGCAGGTTG CAGGCTACAT GAGAAGGTGC AAGTTCAATG 120
CCAAGCAGGG TCCTAGTATT TAATAAAAGC TCAATAAATA TTCATTTTCT TCTTTCCTTC 180

TCTTACTTGA AGTATAACAT TTGATAATGA ATTTTCTCAT TGCAACAATA ACACCCCTTC	240
CACTGAGGGA TTTGTATCCC TGCTTAAGAA GCTATTAGTA TTCTACAGCA GGACTCACCC	300
CACACAATCT TGGCAGGAAT ACATCCCTCT ACCTCTCTGG TCAATAACCT GCCTGGCCTG	360
TGACCCCAAG CTTCTGGAG AAGCACCAAG TCCTCCCAAG TTCCCCC	407

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 267 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CATTGGTGCA GCAGGTTTAG ATGGCTATGT GCTAGAGTAT TGCTTTGAAG GAAGTAAGTA	60
CAACCAGTAG ATAAAATGAA TACTGTCATC AATAGGTGAG ATATGTCCCT CCCCTTTCTG	120
TTGTCTCTCT TTCTTGAGAA CGCATCACCT TCCTACGAAA ATAAGATCAA GCCAAACGTC	180
ATCCTTCTGA GATGTATATA AACTAAGCCC TTTTTTAGTA CTTGGTGCTT ATAAATTGAT	240
ATCTCAAAAG TATCTTGGCT AGGCTGC	267

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 333 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CATAGTCCAG GAGCAGAGTT AGCCAGAATT GCCTCCTGCT GCCCCAGCTT AGAGAGCTCC	60
CATCTCAATC ATTGAGCCTG AAGGCTTCAA GCCCAAATG CAACAAGACC CCCAGCCTAC	120
ATTTCTCAGC TCCCCTGGAG CCAGTGATCC TGTAACGCTG CTGGAGGTCA GTCTGAGCTA	180
CCAAGACTGT CCCTAGACAA AGGTGGGAGT CCCCCACACT GCCAAGACCA AATCCCTCAC	240
TCAACCTGCT GAGGTGTTGG ATGGGGAAAC AAGAGGCAAA ACTGAGGCAC CTGATGCATT	300
CAGCCCTGCT TGTGCAGAAG TGCATTGACT GCC	333

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCTGTGGCGT AAGGCATCCC A

21

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCAAGCACTC CTTTGTA AAA TGTCC

25

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TGCGTTCACC ATTCATGTGG ATGAAGCAG

29

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTCCTACTTC AACTAACCAG TCCACGAG

28

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GATGCTTTGA AGTTATCTCT CTTGG

25

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATCAGTGTGG CAGATATAAT GGACC

25

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GCCCCAAATG CCAGGCTGCA CTGAT

25

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCCAGAAGAC AAGAGTGTGA GCCTT

25

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GCTTCAGGGT GGTCCAATTA GAGTT

25

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TCCAACAACG ACACATTCAG GAGTT

25

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGACACAGAG TAAGATACCC ACTGA

25

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CCTCGGTCTT TGGTCTTTGC ATATC

25

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ACAAGGAAAG TGTCCCTATC TCTGA

25

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CTCGAGGTCT CCCACTGAAG TGCTC

25

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CACTGCACAT TAAGATGGAG CCCGA

25

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCTGTAGAAG TTCTGCTGCG TGTGG

25

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CGAGCTGCCT GACGGCCAGG TCATC

25

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GAAGCATTTG CGGTGGACGA TGGAG

25

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TAGAAGACCA AATGCCCCGA GT

22

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TGTATTTCTG TGGGATCGGT GG

22

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 369 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```

CCATAAGAGA AATGATTGGT AGGTTTGCAT GAAATTTTAA AATTTCTGT GGC GTAAGGC      60
ATCCCATAAC GAAGCCAAAA GGTGAGTGAT AGACTGGGAG AAATAACTGC CAGACGTTGC      120
CAGACAAAGA TTTCATATTT CTAATATGCT AGAGTACCTT TAATTTGATA AGAAAAAGAT      180
AAGCAATCCT GTAATAAAAT GGACATTTTA CAAAGGAGTG CTTGCAAATG GCCAGTGAAT      240
TTATGCAAAT ATGTT CAGGG AAATAGGAAT GAAAACGAGA TTCCACTTTT TCATCATCCA      300
TTTGATTGGC AAGAAATTTT TAAAAGAGTA ATACCTAGTG AATCACTCAT GTAGGAAAAT      360
GGGTTGGTG                                     369

```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 301 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 212
- (D) OTHER INFORMATION: /note= "N = A, C, G, or T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```

GCCCTTGAAG AGTGTAACCA AGAAGCATCT CTCAATCAAT GAACCTGAGA CAGCCTGTTC      60
ACTTCTGACC ATCATTCTTG TCCTTTAGAT CTCAGTTTCA AATTCATTTC TTCTAGACAT      120
TCATCTCTTC CCATGTTTAA TCTGGAACCA TCTACCTTC CACCAGACCA ATTATCCTGG      180
CAAATTAATG TAATAGACCA GTATTAATTA TNTGGTTGTA TGTCTTAACA ACATTCTAGG      240
TGCTGTGCCA AAAACAAATG AATAGCAACA CAAGGTCTTC TTGGTTACAC TCTTCAAGGG      300
C                                     301

```

SECRET

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3061 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 15..1172

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CGGCTCTCCT CAAC ATG AGA GCT GCA CCC CTC CTC CTG GCC AGG GCA GCA 50
Met Arg Ala Ala Pro Leu Leu Leu Ala Arg Ala Ala
1 5 10

AGC CTT AGC CTT GGC TTC TTG TTT CTG CTT TTT TTC TGG CTA GAC CGA 98
Ser Leu Ser Leu Gly Phe Leu Phe Leu Leu Phe Phe Trp Leu Asp Arg
15 20 25

AGT GTA CTA GCC AAG GAG TTG AAG TTT GTG ACT TTG GTG TTT CGG CAT 146
Ser Val Leu Ala Lys Glu Leu Lys Phe Val Thr Leu Val Phe Arg His
30 35 40

GGA	GAC	CGA	AGT	CCC	ATT	GAC	ACC	TTT	CCC	ACT	GAC	CCC	ATA	AAG	GAA	194
Gly	Asp	Arg	Ser	Pro	Ile	Asp	Thr	Phe	Pro	Thr	Asp	Pro	Ile	Lys	Glu	
45					50					55					60	

TCC	TCA	TGG	CCA	CAA	GGA	TTT	GGC	CAA	CTC	ACC	CAG	CTG	GGC	ATG	GAG	242
Ser	Ser	Trp	Pro	Gln	Gly	Phe	Gly	Gln	Leu	Thr	Gln	Leu	Gly	Met	Glu	
				65					70					75		

CAG CAT TAT GAA CTT GGA GAG TAT ATA AGA AAG AGA TAT AGA AAA TTC 290
Gln His Tyr Glu Leu Gly Glu Tyr Ile Arg Lys Arg Tyr Arg Lys Phe
80 85 90

TTG AAT GAG TCC TAT AAA CAT GAA CAG GTT TAT ATT CGA AGC ACA GAC 338
Leu Asn Glu Ser Tyr Lys His Glu Gln Val Tyr Ile Arg Ser Thr Asp
95 100 105

GTT GAC CGG ACT TTG ATG AGT GCT ATG ACA AAC CTG GCA GCC CTG TTT 386
Val Asp Arg Thr Leu Met Ser Ala Met Thr Asn Leu Ala Ala Leu Phe
110 115 120

CCC CCA GAA GGT GTC AGC ATC TGG AAT CCT ATC CTA CTC TGG CAG CCC 434
Pro Pro Glu Gly Val Ser Ile Trp Asn Pro Ile Leu Leu Trp Gln Pro
125 130 135 140

ATC CCG GTG CAC ACA GTT CCT CTT TCT GAA GAT CAG TTG CTA TAC CTG 482
Ile Pro Val His Thr Val Pro Leu Ser Glu Asp Gln Leu Leu Tyr Leu

	145		150		155	
CCT TTC AGG AAC TGC CCT CGT TTT CAA GAA CTT GAG AGT GAG ACT TTG						530
Pro Phe Arg Asn Cys Pro Arg Phe Gln Glu Leu Glu Ser Glu Thr Leu						
	160		165		170	
AAA TCA GAG GAA TTC CAG AAG AGG CTG CAC CCT TAT AAG GAT TTT ATA						578
Lys Ser Glu Glu Phe Gln Lys Arg Leu His Pro Tyr Lys Asp Phe Ile						
	175		180		185	
GCT ACC TTG GGA AAA CTT TCA GGA TTA CAT GGC CAG GAC CTT TTT GGA						626
Ala Thr Leu Gly Lys Leu Ser Gly Leu His Gly Gln Asp Leu Phe Gly						
	190		195		200	
ATT TGG AGT AAA GTC TAC GAC CCT TTA TAT TGT GAG AGT GTT CAC AAT						674
Ile Trp Ser Lys Val Tyr Asp Pro Leu Tyr Cys Glu Ser Val His Asn						
205		210		215		220
TTC ACT TTA CCC TCC TGG GCC ACT GAG GAC ACC ATG ACT AAG TTG AGA						722
Phe Thr Leu Pro Ser Trp Ala Thr Glu Asp Thr Met Thr Lys Leu Arg						
	225		230		235	
GAA TTG TCA GAA TTG TCC CTC CTG TCC CTC TAT GGA ATT CAC AAG CAG						770
Glu Leu Ser Glu Leu Ser Leu Leu Ser Leu Tyr Gly Ile His Lys Gln						
	240		245		250	
AAA GAG AAA TCT AGG CTC CAA GGG GGT GTC CTG GTC AAT GAA ATC CTC						818
Lys Glu Lys Ser Arg Leu Gln Gly Gly Val Leu Val Asn Glu Ile Leu						
	255		260		265	
AAT CAC ATG AAG AGA GCA ACT CAG ATA CCA AGC TAC AAA AAA CTT ATC						866
Asn His Met Lys Arg Ala Thr Gln Ile Pro Ser Tyr Lys Lys Leu Ile						
	270		275		280	
ATG TAT TCT GCG CAT GAC ACT ACT GTG AGT GGC CTA CAG ATG GCG CTA						914
Met Tyr Ser Ala His Asp Thr Thr Val Ser Gly Leu Gln Met Ala Leu						
285		290		295		300
GAT GTT TAC AAC GGA CTC CTT CCT CCC TAT GCT TCT TGC CAC TTG ACG						962
Asp Val Tyr Asn Gly Leu Leu Pro Pro Tyr Ala Ser Cys His Leu Thr						
	305		310		315	
GAA TTG TAC TTT GAG AAG GGG GAG TAC TTT GTG GAG ATG TAC TAT CGG						1010
Glu Leu Tyr Phe Glu Lys Gly Glu Tyr Phe Val Glu Met Tyr Tyr Arg						
	320		325		330	
AAT GAG ACG CAG CAC GAG CCG TAT CCC CTC ATG CTA CCT GGC TGC AGC						1058
Asn Glu Thr Gln His Glu Pro Tyr Pro Leu Met Leu Pro Gly Cys Ser						
	335		340		345	
CCT AGC TGT CCT CTG GAG AGG TTT GCT GAG CTG GTT GGC CCT GTG ATC						1106
Pro Ser Cys Pro Leu Glu Arg Phe Ala Glu Leu Val Gly Pro Val Ile						
	350		355		360	

CCT CAA GAC TGG TCC ACG GAG TGT ATG ACC ACA AAC AGC CAT CAA GGT	1154
Pro Gln Asp Trp Ser Thr Glu Cys Met Thr Thr Asn Ser His Gln Gly	
365 370 375 380	
ACT GAG GAC AGT ACA GAT TAGTGTGCAC AGAGATCTCT GTAGAAAGAG	1202
Thr Glu Asp Ser Thr Asp	
385	
TAGCTGCCCT TTCTCAGGGC AGATGATGCT TTGAGAACAT ACTTTGGCCA TTACCCCCCA	1262
GCTTTGAGGA AAATGGGCTT TGGATGATTA TTTTATGTTT TAGGGACCCC CAACCTCAGG	1322
CAATTCCTAC CTCTTCACCT GACCCTGCCC CCACTTGCCA TAAACCTTAG CTAAGTTTTG	1382
TTTTGTTTTT CAGCGTTAAT GTAAAGGGGC AGCAGTGCCA AAATATAATC AGAGATAAAG	1442
CTTAGGTCAA AGTTCATAGA GTTCCCATGA ACTATATGAC TGGCCACACA GGATCTTTTG	1502
TATTTAAGGA TTCTGAGATT TTGCTTGAGC AGGATTAGAT AAGTCTGTTC TTTAAATTTT	1562
TGAAATGGAA CAGATTTCAA AAAAAATTCC CACAATCTAG GGTGGGAACA AGGAAGGAAA	1622
GATGTGAATA GGCTGATGGG GAAAAAACCA ATTTACCCAT CAGTTCCAGC CTTCTCTCAA	1682
GGAGAGGCAA AGAAAGGAGA TACAGTGAG ACATCTGGAA AGTTTTCTCC ACTGGAAAAC	1742
TGCTACTATC TGTTTTTATA TTTCTGTAA AATATATGAG GCTACAGAAC TAAAAATTAA	1802
AACCTCTTTG TGTCCCTTGG TCCTGGAACA TTTATGTTCC TTTTAAAGAA ACAAAAAATCA	1862
AACTTTACAG AAAGATTTGA TGTATGTAAT ACATATAGCA GCTCTTGAAG TATATATATC	1922
ATAGCAAATA AGTCATCTGA TGAGAACAAAG CTATTTGGGC ACAACACATC AGGAAAGAGA	1982
GCACCACGTG ATGGAGTTTC TCCAGAAGCT CCAGTGATAA GAGATGTTGA CTCTAAAGTT	2042
GATTTAAGGC CAGGCATGGT GGTTTACGCC TATAATCCCA GCATTTTGGG ACTCCGAGGT	2102
GGGCAGATCA CTTGAGCTCA GGAGCTCAAG ATCAGCCTGG GCAACATGGT GAAACCTTGT	2162
CTCTACATAA AATACAAAAA CTTAGATGGG CATGGTGCTG TGTGCCTATA GTCCACTACT	2222
TGTGGGGCTA AGGCAGGAGG ATCACTTGAG CCCCAGAGGT CGAGGCTACA GTGACCCAAG	2282
AGTGCACTAC TGTACTCCAG CCAGGGCAAG AGAGCGAGAC CCTGTCTCAA TAAATAAATA	2342
AATAAATAAA TAAATAAATA AATAAAAACA AAGTTGATTA AGAAAGGAAG TATAGGCCAG	2402
GCACAGTGGC TCACACCTGT AATCCTTGCA TTTTGGAAGG CTGAGGCAGG AGGATCACTT	2462
TAGGCCTGGT GTGTTCAAGA CCAGCCTGGT CAACATAGTG AGACACTGTC TCTACCAAAA	2522
AAAGGAAGGA AGGGACACAT ATCAAAGTGA AACAAAATTA GAAATGTAAT TATGTTATGT	2582

TCTAAGTGCC TCCAAGTTCA AAACCTTATTG GAATGTTGAG AGTGTGGTTA CGAAATACGT 2642
TAGGAGGACA AAAGGAATGT GTAAGTCTTT AATGCCGATA TCTTCAGAAA ACCTAAGCAA 2702
ACTTACAGGT CCTGCTGAAA CTGCCCCACTC TGCAAGAAGA AATCATGATA TAGCTTTCCA 2762
TGTGGCAGAT CTACATGTCT AGAGAACACT GTGCTCTATT ACCATTATGG ATAAAGATGA 2822
GATGGTTTCT AGAGATGGTT TCTACTGGCT GCCAGAATCT AGAGCAAAGC CATCCCCCCT 2882
CCTGGTTGGT CACAGAATGA CTGACAAAGA CATCGATTGA TATGCTTCTT TGTGTTATTT 2942
CCCTCCCAAG TAAATGTTTG TCCTTGGGTC CATTTTCTAT GCTTGTAACGT GTCTTCTAGC 3002
AGTGAGCCAA ATGTAAAATA GTGAATAAAG TCATTATTAG GAAGTTCAAA AAAAAAAAAA 3061

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 386 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met	Arg	Ala	Ala	Pro	Leu	Leu	Leu	Ala	Arg	Ala	Ala	Ser	Leu	Ser	Leu	1	5	10	15
Gly	Phe	Leu	Phe	Leu	Leu	Phe	Phe	Trp	Leu	Asp	Arg	Ser	Val	Leu	Ala	20	25	30	
Lys	Glu	Leu	Lys	Phe	Val	Thr	Leu	Val	Phe	Arg	His	Gly	Asp	Arg	Ser	35	40	45	
Pro	Ile	Asp	Thr	Phe	Pro	Thr	Asp	Pro	Ile	Lys	Glu	Ser	Ser	Trp	Pro	50	55	60	
Gln	Gly	Phe	Gly	Gln	Leu	Thr	Gln	Leu	Gly	Met	Glu	Gln	His	Tyr	Glu	65	70	75	80
Leu	Gly	Glu	Tyr	Ile	Arg	Lys	Arg	Tyr	Arg	Lys	Phe	Leu	Asn	Glu	Ser	85	90	95	
Tyr	Lys	His	Glu	Gln	Val	Tyr	Ile	Arg	Ser	Thr	Asp	Val	Asp	Arg	Thr	100	105	110	
Leu	Met	Ser	Ala	Met	Thr	Asn	Leu	Ala	Ala	Leu	Phe	Pro	Pro	Glu	Gly	115	120	125	
Val	Ser	Ile	Trp	Asn	Pro	Ile	Leu	Leu	Trp	Gln	Pro	Ile	Pro	Val	His	130	135	140	

Thr Val Pro Leu Ser Glu Asp Gln Leu Leu Tyr Leu Pro Phe Arg Asn
 145 150 155 160
 Cys Pro Arg Phe Gln Glu Leu Glu Ser Glu Thr Leu Lys Ser Glu Glu
 165 170 175
 Phe Gln Lys Arg Leu His Pro Tyr Lys Asp Phe Ile Ala Thr Leu Gly
 180 185 190
 Lys Leu Ser Gly Leu His Gly Gln Asp Leu Phe Gly Ile Trp Ser Lys
 195 200 205
 Val Tyr Asp Pro Leu Tyr Cys Glu Ser Val His Asn Phe Thr Leu Pro
 210 215 220
 Ser Trp Ala Thr Glu Asp Thr Met Thr Lys Leu Arg Glu Leu Ser Glu
 225 230 235 240
 Leu Ser Leu Leu Ser Leu Tyr Gly Ile His Lys Gln Lys Glu Lys Ser
 245 250 255
 Arg Leu Gln Gly Gly Val Leu Val Asn Glu Ile Leu Asn His Met Lys
 260 265 270
 Arg Ala Thr Gln Ile Pro Ser Tyr Lys Lys Leu Ile Met Tyr Ser Ala
 275 280 285
 His Asp Thr Thr Val Ser Gly Leu Gln Met Ala Leu Asp Val Tyr Asn
 290 295 300
 Gly Leu Leu Pro Pro Tyr Ala Ser Cys His Leu Thr Glu Leu Tyr Phe
 305 310 315 320
 Glu Lys Gly Glu Tyr Phe Val Glu Met Tyr Tyr Arg Asn Glu Thr Gln
 325 330 335
 His Glu Pro Tyr Pro Leu Met Leu Pro Gly Cys Ser Pro Ser Cys Pro
 340 345 350
 Leu Glu Arg Phe Ala Glu Leu Val Gly Pro Val Ile Pro Gln Asp Trp
 355 360 365
 Ser Thr Glu Cys Met Thr Thr Asn Ser His Gln Gly Thr Glu Asp Ser
 370 375 380
 Thr Asp
 385

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

- [illegible]

TCGCTCCACA TTCATCCTTT CT

(2) INFORMATION FOR SEQ ID NO:50:

- (xi) SEQUENCE DESCRIPTION: SEO ID NO:50:

25

(2) INFORMATION FOR SEQ ID NO:51:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

25

(2) INFORMATION FOR SEO ID NO:52:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

25

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CACGATGCCA TTCTGCCATT TCTGT

25

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GGAAGAGATG GAATAGAAAC TGTA

25

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CTTAAC TCGG GCATT TGGTC TTC

23

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Arg Lys Lys Glu Lys Val Lys Arg Ser Gln Lys Ala Thr Glu Phe Ile
1 5 10 15

Asp Tyr Ser Ile Glu
20

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

AAACAAACGT CTTTGGGTAA A

21

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CTGGACAAAG AGGAATATGA

20

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GCCCTTTATA AATACGATTA GTATGGAG

28

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TGTAGTTAGT GCAGCAAAAG GAAGA

25

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GATGTAATTA AAGCTGTAGA TGAGGG

26

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GAATACTAAC AATCTGCTCA AACTTGGG

28

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GCCAAATGGG TAGCATTGTT GCTCGG

26

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CAGAGTGGGG CAAGATACCC TTGAG

25

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

AATGGAATTT CTTATGCCCT C

21

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CAATGCCAAG CACCCACTGA TTC

23

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ACACAGACAC ACACATGCAC ACCA

24

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CCTACCTGTG CAGAAATCAA

20

(2) INFORMATION FOR SEQ ID NO:73:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

AGCAGCATAG CCTCTCTGAA ACTC

24

(2) INFORMATION FOR SEQ ID NO:74:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CCTTCTCATG TAGCCTGCAA CCTGCTC

27

(2) INFORMATION FOR SEQ ID NO:75:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CATTGGTGCA GCAGGTTTAG ATGG

24

(2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GAGATATCAA TTTATAAGCA CCAAG

25

(2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

ATCTCAATCA TTGAGCCTGA AGG

23

(2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CAGCAGGTTG AGTGAGGGAT TTGG

24

(2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

CGCCTCAGGC TGGGGCAGCA TT

22

(2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

ACAGTGGAAG AGTCTCATTC GAGAT

25

(2) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

CGAGCTGCCT GACGGCCAGG TCATC

25

(2) INFORMATION FOR SEQ ID NO:82:

- (i) SEQUENCE CHARACTERISTICS:

[illegible]

GAAGCATTG CGGTGGACGA TGGAG

(2) INFORMATION FOR SEO ID NO:83:

(A) LENGTH: 2087 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) NAME/KEY: CDS
(B) LOCATION: 99..503

GACCTTAAAT ATATCGAGGT GGCTAATTGA TGTATAATAA TTTACAAAAT TATTCTTCTA 60

AGG AAC CAG AAA TAT GAG GAT ATG CAC AAT ATT ATT CAC ATT TTA CAG 161
Arg Asn Gln Lys Tyr Glu Asp Met His Asn Ile Ile His Ile Leu Gln
10 15 20

ATC AGA AAA TTG AGG CAC AGA TTA AGT AAC TTC CCA AGG CTA CCA GGC 209
Ile Arg Lys Leu Arg His Arg Leu Ser Asn Phe Pro Arg Leu Pro Gly
25 30 35

ATT CTA GCT CCA GAA ACT GTG CTC TTA CCA TTC TGC TAC AAG GTA TTT 257
Ile Leu Ala Pro Glu Thr Val Leu Leu Pro Phe Cys Tyr Lys Val Phe
40 45 50

CGA AAA AAA GAA AAA GTA AAA AGA AGT CAA AAG GCA ACA GAG TTC ATT 305
Arg Lys Lys Glu Lys Val Lys Arg Ser Gln Lys Ala Thr Glu Phe Ile
55 60 65

GAT TAT TCC ATA GAA CAG TCA CAC CAT GCA ATT CTC ACA CCC TTG CAG 353
Asp Tyr Ser Ile Glu Gln Ser His His Ala Ile Leu Thr Pro Leu Gln
70 75 80 85

ACA	CAC	TTG	ACC	ATG	AAA	GGT	TCC	TCA	ATG	AAA	TGT	TCC	TCA	TTA	TCT	401
Thr	His	Leu	Thr	Met	Lys	Gly	Ser	Ser	Met	Lys	Cys	Ser	Ser	Leu	Ser	
				90					95					100		

TCA GAA GCC ATA TTA TTA ACA TTG ACT TTG CAG TTA ACT CAG ACC CTA	449
Ser Glu Ala Ile Leu Phe Thr Leu Thr Leu Gln Leu Thr Gln Thr Leu	
105 110 115	
GGT CTG GAA TGC TGT CTT CTC TAC TTA TCC AAA ACT ATA CAT CCA CAG	497
Gly Leu Glu Cys Cys Leu Leu Tyr Leu Ser Lys Thr Ile His Pro Gln	
120 125 130	
ATC ATA TAAACTCTCA GCCCTGCTGC AAAGCCTTTC CAGAAAAATA AAAATGGTTG	553
Ile Ile	
135	
AAAAGGCAAT TCTGCTACCA ATGACTGTTT AAGCCCAGCC AAGTAACTGA ACCATTCCAA	613
CTTCAATTTA CTTATGAAAA GAATTTGATG ATGTAGGAGG TTATTTCAAT TCTAAAATAC	673
AAACCCATGT TGATCTTTCT CAATCTTGAA CTCATAGATT ATTATCTATT ATCTCAATTT	733
AGTTTGTTAT TTATCCTAGT GGGCCATTAA AACTACCAC ATGTGTTTCT GTCTCTCCAT	793
TAGTCAATAA CTAAACTAAC GAGCAATTAG TAAGCCATGT GCCAGATGCT CCGCTAGGCA	853
CCAGAGGGAT AAAAAACAATA CTTATAGTAT ACCACTAATT TTCGCTTAGT AACTAGTGAA	913
ATGTTCAAGT CATGCCTGAG TCAAGAGTTG AGGAGACATT ACAATGTGTA ATGGAAACCA	973
AGGAAAGTGA AACTTTGGAT AAGTGGGGAC TAGTGTATTT ATATATTTAA TTGATTTCTG	1033
ACTCTATCAT TGGCCTCCAA ACACAGATTG TGTTTTTCTT TGGTTTTGTT TTCTTCACTA	1093
TGGGATCTTC TGTGCCCAGC ACAGTGCCTG ACACATAGAA AACAATCAAT ATTTGCTGAA	1153
TAAATGATTA AAAAATCAGA GAACTTTCCC ATTCTGTTTG GATCTATAGA ACATCCAGAG	1213
TAAGTGATGA GGGCCTCTGC ATTTATATGC GCTTAAATTA AGATTATGTG AGAAAAGTTT	1273
AAAGACACTT AGTAGAGTGA TTTTGAAATA TAGTAAACAC TTGGAAATGG TGGTGCTTTA	1333
AAAAGATATT AATAGATAAT ATGAAAATCT CCATCTCAAA AATAATGCAT AACTATTTA	1393
AAGGAAAATC ACATCTCCAG GCTTTCAATG TTTGTTTCATT ACTTTTTTCAT ATATTTTTAC	1453
CATCTGCTGA AGGCAGTCAT ATCAAAGGGT AAAGAAAGAT GGGAGGAAAA CTCAGTAAGA	1513
ATTATATTAG TCTGTTTGCA AAGTAGAAAA AGATTCTCAT CACTCAACCT TATGAGCAGG	1573
AAGAGGGAAG GCTGTTTGAG AACCATTTAC TTAGCAGAAC CACATATTTT AGACACTTCC	1633
CTGCATTAAC TGCACAAACA ATATGTTTGC AAACCTGTTG ATCAACCTCC AACACGACA	1693
CATTCAGGAG TTAAATATTT TTCATCAAAC ATTGGATTTT TCCTTAACGC TAGAGATTGC	1753
TACAAATCTT CTGAAGGGTC TCAATGGCTT CAGGCTAAGA AGAGATTTCT CCCTGTTATA	1813

AGCAGCAAGA CAAATTAGCC ATTTCACTCT CAACTTCAC TAATGATCAC ATTCTTTCCA 1873
AAAGGAACTC TAGAAGACCA AATGCCCCGA GTTAAGAACA TCAAAACTAA CCATCTGAAG 1933
AAACTTCCCA AGTGTAAGAC TCTGCCTGCA CGACAACACA TAAAAAAGA GAGAAGAATC 1993
AAATAGACAC AATAAAAAAT GATAAAGGGG ATATCACCAC CGATCCCACA GAAATACAAA 2053
CTACCATCAG AGAATACTAC AAACACCTCT ACGC 2087

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Met Arg Ala Phe Leu Arg Asn Gln Lys Tyr Glu Asp Met His Asn Ile
1 5 10 15
Ile His Ile Leu Gln Ile Arg Lys Leu Arg His Arg Leu Ser Asn Phe
20 25 30
Pro Arg Leu Pro Gly Ile Leu Ala Pro Glu Thr Val Leu Leu Pro Phe
35 40 45
Cys Tyr Lys Val Phe Arg Lys Lys Glu Lys Val Lys Arg Ser Gln Lys
50 55 60
Ala Thr Glu Phe Ile Asp Tyr Ser Ile Glu Gln Ser His His Ala Ile
65 70 75 80
Leu Thr Pro Leu Gln Thr His Leu Thr Met Lys Gly Ser Ser Met Lys
85 90 95
Cys Ser Ser Leu Ser Ser Glu Ala Ile Leu Phe Thr Leu Thr Leu Gln
100 105 110
Leu Thr Gln Thr Leu Gly Leu Glu Cys Cys Leu Leu Tyr Leu Ser Lys
115 120 125
Thr Ile His Pro Gln Ile Ile
130 135

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2505 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 99..503

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

GACCTTAAAT ATATCGAGGT GGCTAATTGA TGTATAATAA TTTACAAAAT TATTCTTCTA	60
TTGCTACAGA GCTACAATTC AATTTACAGT AGGCCACC ATG AGG GCC TTC TTA	113
Met Arg Ala Phe Leu	
1 5	
AGG AAC CAG AAA TAT GAG GAT ATG CAC AAT ATT ATT CAC ATT TTA CAG	161
Arg Asn Gln Lys Tyr Glu Asp Met His Asn Ile Ile His Ile Leu Gln	
10 15 20	
ATC AGA AAA TTG AGG CAC AGA TTA AGT AAC TTC CCA AGG CTA CCA GGC	209
Ile Arg Lys Leu Arg His Arg Leu Ser Asn Phe Pro Arg Leu Pro Gly	
25 30 35	
ATT CTA GCT CCA GAA ACT GTG CTC TTA CCA TTC TGC TAC AAG GTA TTT	257
Ile Leu Ala Pro Glu Thr Val Leu Leu Pro Phe Cys Tyr Lys Val Phe	
40 45 50	
CGA AAA AAA GAA AAA GTA AAA AGA AGT CAA AAG GCA ACA GAG TTC ATT	305
Arg Lys Lys Glu Lys Val Lys Arg Ser Gln Lys Ala Thr Glu Phe Ile	
55 60 65	
GAT TAT TCC ATA GAA CAG TCA CAC CAT GCA ATT CTC ACA CCC TTG CAG	353
Asp Tyr Ser Ile Glu Gln Ser His His Ala Ile Leu Thr Pro Leu Gln	
70 75 80 85	
ACA CAC TTG ACC ATG AAA GGT TCC TCA ATG AAA TGT TCC TCA TTA TCT	401
Thr His Leu Thr Met Lys Gly Ser Ser Met Lys Cys Ser Ser Leu Ser	
90 95 100	
TCA GAA GCC ATA TTA TTC ACA TTG ACT TTG CAG TTA ACT CAG ACC CTA	449
Ser Glu Ala Ile Leu Phe Thr Leu Thr Leu Gln Leu Thr Gln Thr Leu	
105 110 115	
GGT CTG GAA TGC TGT CTT CTC TAC TTA TCC AAA ACT ATA CAT CCA CAG	497
Gly Leu Glu Cys Cys Leu Leu Tyr Leu Ser Lys Thr Ile His Pro Gln	
120 125 130	
ATC ATA TAACTCTCA GCCCTGCTGC AAAGCCTTTC CAGAAAAATA AAAATGGTTG	553
Ile Ile	
135	
AAAAGGCAAT TCTGCTACCA ATGACTGTTT AAGCCCAGCC AAGTAACTGA ACCATTCCAA	613

CTTCAATTTA CTTATGAAAA GAATTTTGATG ATGTAGGAGG TTATTTCAAT TCTAAAATAC	673
AAACCCATGT TGATCTTTCT CAATCTTGAA CTCATAGATT ATTATCTATT ATCTCAATTT	733
AGTTTGTTAT TTATCCTAGT GGGCCATTAA AAACCTACCAC ATGTGTTTCT GTCTCTCCAT	793
TAGTCAATAA CTAAACTAAC GAGCAATTAG TAAGCCATGT GCCAGATGCT CCGCTAGGCA	853
CCAGAGGGAT AAAAAACAATA CTTATAGTAT ACCACTAATT TTCGCTTAGT AACTAGTGAA	913
ATGTTCAAGT CATGCCTGAG TCAAGAGTTG AGGAGACATT ACAATGTGTA ATGGAAACCA	973
AGGAAAGTGA AACTTTGGAT AAGTGGGGAC TAGTGTATTT ATATATTTAA TTGATTTCTG	1033
ACTCTATCAT TGGCCTCCAA ACACAGATTG TGTTTTTCTT TGGTTTTGTT TTCTTCACTA	1093
TGGGATCTTC TGTGCCCAGC ACAGTGCCTG ACACATAGAA AACAATCAAT ATTTGCTGAA	1153
TAAATGATTA AAAAATCAGA GAACTTTCCC ATTCTGTTTG GATCTATAGA ACATCCAGAG	1213
TAAGTGATGA GGGCCTCTGC ATTTATATGC GCTTAAATTA AGATTATGTG AGAAAAGTTT	1273
AAAGACACTT AGTAGAGTGA TTTTGAAATA TAGTAAACAC TTGGAAATGG TGGTGCTTTA	1333
AAAAGATATT AATAGATAAT ATGAAAATCT CCATCTCAAA AATAATGCAT AAACATTTTA	1393
AAGGAAAATC ACATCTCCAG GCTTCAATG TTTGTTCATT ACTTTTTCAT ATATTTTAC	1453
CATCTGCTGA AGGCAGTCAT ATCAAAGGGT AAAGAAAGAT GGGAGGAAAA CTCAGTAAGA	1513
ATTATATTAG TCTGTTTGCA AAGTAGAAAA AGATTCTCAT CACTCAACCT TATGAGCAGG	1573
AAGAGGGAAG GCTGTTTGAG AACCATTAC TTAGCAGAAC CACATATTTT AGACACTTCC	1633
CTGCATTAAC TGCACAAACA ATATGTTTGC AAACCTGTTG ATCAACCTCC AACAACGACA	1693
CATTCAGGAG TTAAATATTT TTCATCAAAC ATTGGATTTT TCCTTAACGC TAGAGATTGC	1753
TACAAATCTT CTGAAGGGTC TCAATGGCTT CAGGCTAAGA AGAGATTTCT CCCTGTTATA	1813
AGCAGCAAGA CAAATTAGCC ATTTCACTCT CAACTTCAC TAATGATCAC ATTCTTTCCA	1873
AAAGGAACTC TAGAAGACCA AATGCCCCGA GTTAAGAACA TCAAACTAA CCATCTGAAG	1933
AAACTTCCCA AGTGTAAGAC TCTGCCATTA AACATTACC GAGAGGGGAC TCAAACAGTC	1993
TTTCTTCCTT TGTCGTGTTT CTTGCTCCCA GACCAAGGCA CTGACGACAG TACTGATACA	2053
TAATTTAAAA GCACACTCCC TTCCACTTTG GTAATACCAG AACTCTAATT GGACCACCCT	2113
GAAGCTTAGG ACTACCAGCC ATACAAATAG TAAACTCTGT CCACGATTCA CTCATCTGTG	2173
TATTTTCTAT AGATGTTTAC TAGGCGTTTG TTATATAAAA ATACCCCGGC CAGGCACGGT	2233

GGCTCACGCC TGTAATCCCA GCACTTTGGG AGGTGGGTGG ATCACCTGAG GTCGGGAGTT 2293
 CGAGACCAGC CTGACCAGCA TGGTGAACCC CCCATCTCTA CTAAAAACAC AAAAAATTAG 2353
 CCGGGCGTGG TGGCACATGC CTGTAATCCC AGCTACTCAG GAGGCTGAGG CGGAGAATTG 2413
 CTTGAACCCG GAAGGTGGAG GTTGTTCGGG TGAGCTGAGA TTGCACTATT GCACTCCAGC 2473
 CTGGGCAACA GGAGTAAAC TCCCCCCCAC CC 2505

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 135 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Met	Arg	Ala	Phe	Leu	Arg	Asn	Gln	Lys	Tyr	Glu	Asp	Met	His	Asn	Ile
1				5					10					15	
Ile	His	Ile	Leu	Gln	Ile	Arg	Lys	Leu	Arg	His	Arg	Leu	Ser	Asn	Phe
			20					25						30	
Pro	Arg	Leu	Pro	Gly	Ile	Leu	Ala	Pro	Glu	Thr	Val	Leu	Leu	Pro	Phe
			35				40							45	
Cys	Tyr	Lys	Val	Phe	Arg	Lys	Lys	Glu	Lys	Val	Lys	Arg	Ser	Gln	Lys
			50				55							60	
Ala	Thr	Glu	Phe	Ile	Asp	Tyr	Ser	Ile	Glu	Gln	Ser	His	His	Ala	Ile
			65				70							75	80
Leu	Thr	Pro	Leu	Gln	Thr	His	Leu	Thr	Met	Lys	Gly	Ser	Ser	Met	Lys
				85					90					95	
Cys	Ser	Ser	Leu	Ser	Ser	Glu	Ala	Ile	Leu	Phe	Thr	Leu	Thr	Leu	Gln
			100					105						110	
Leu	Thr	Gln	Thr	Leu	Gly	Leu	Glu	Cys	Cys	Leu	Leu	Tyr	Leu	Ser	Lys
			115				120							125	
Thr	Ile	His	Pro	Gln	Ile	Ile									
			130			135									

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

CCCACCTCCC AAAGTGCTGG GA

22

TOPOT-9454-1001